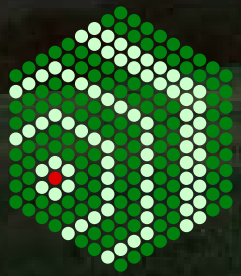
Radiation Damage Induced Phasing

Max Nanao

EMBL Grenoble



EMBL
Grenoble Outstation

Goals

- *De novo* phasing of crystal structures
- But more broadly:
  - *Model radiation damage to improve data quality*
  - *AND get phases*

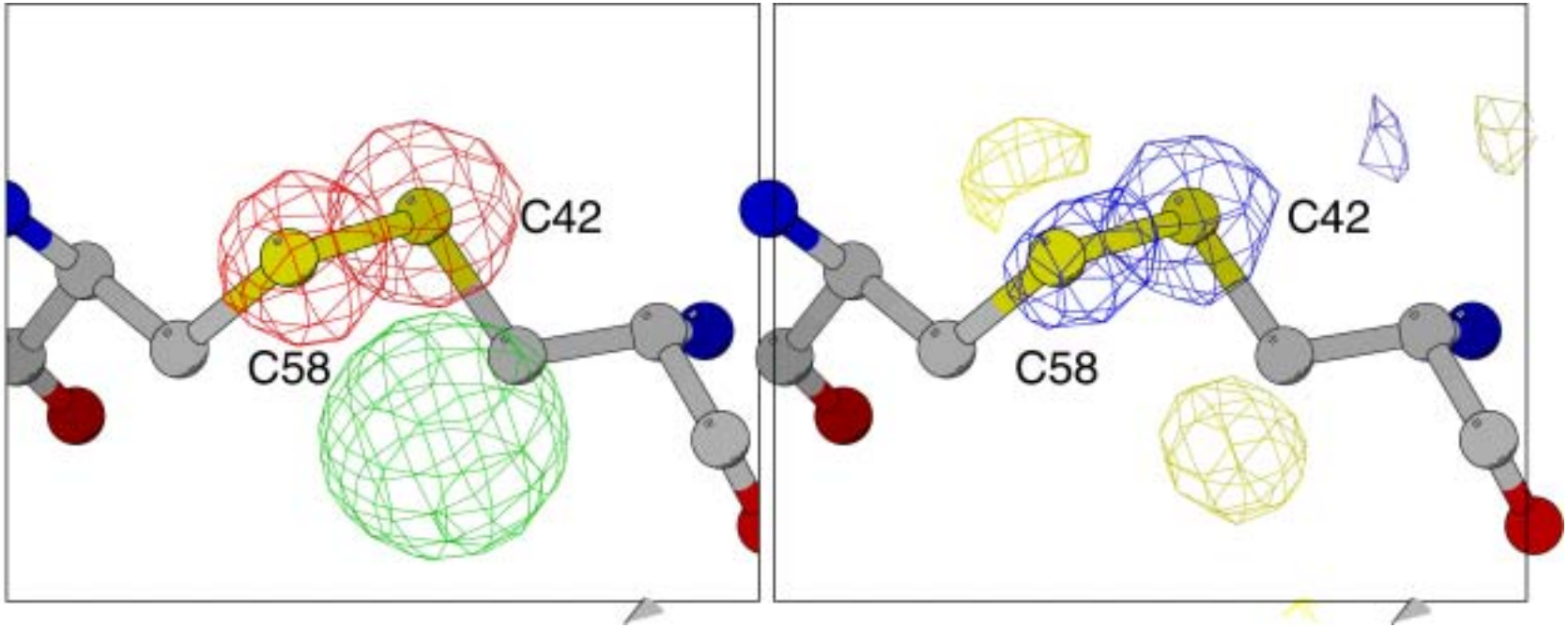
Radiation damage in macromolecular crystals

- Specific effects
  - Lowered occupancy of heavy atoms (Se, Br, Hg…)
  - Breakage of S-S bonds
  - Decarboxylation
  - Debromination of nucleic acids
  - Dehydroxylation

  ==>Signal
- Global effects
  - Increase in unit cell dimensions
  - Small rotations

  ==>Noise

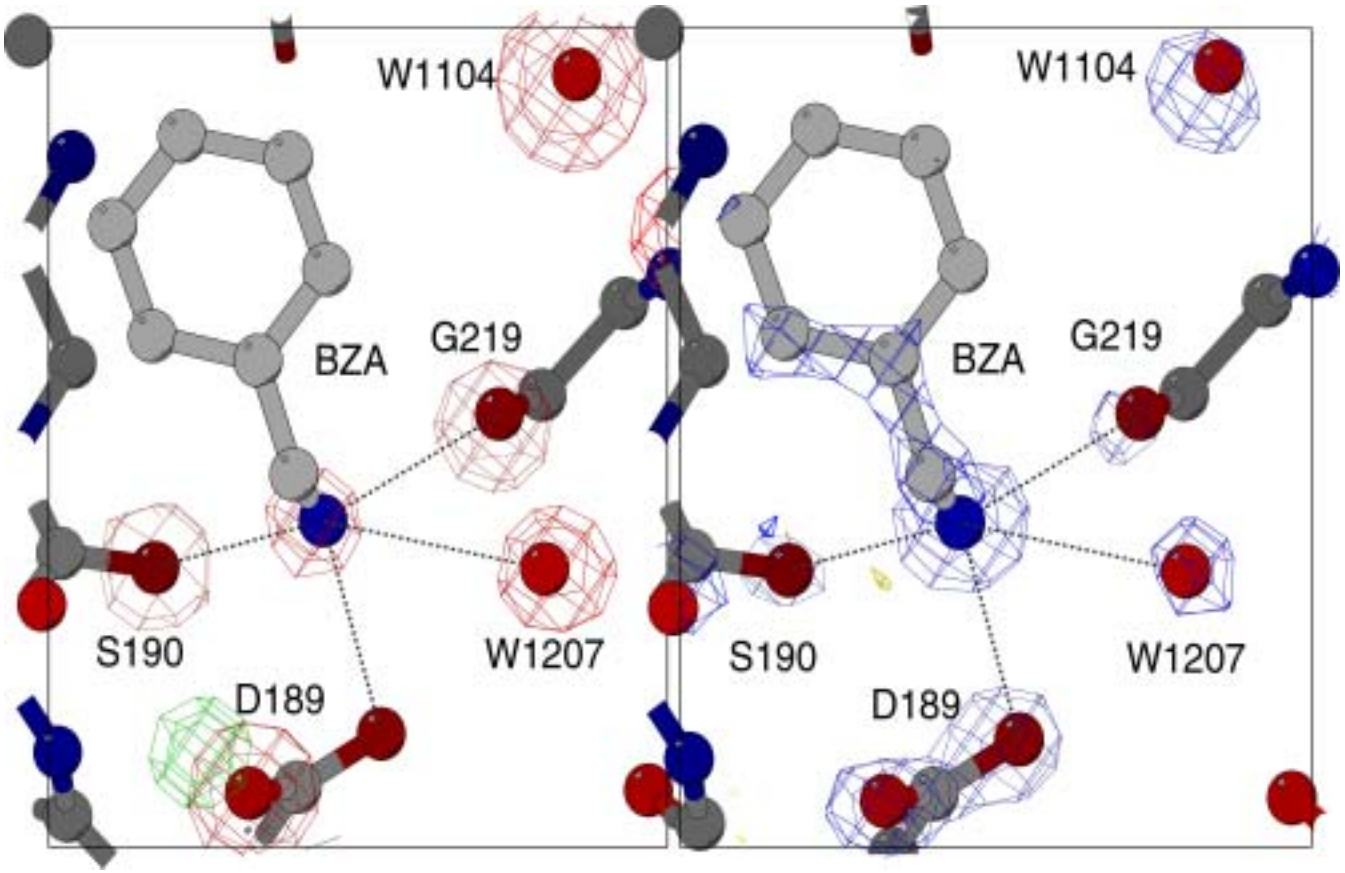
Specific Effects



Red & Green: Sharp sites

Blue & Yellow: Difference maps

More Specific effects



Phasing by Radiation Damage

- These specific sites can be loosely considered similar to SIR
- “Heavy atom” dataset = dataset before radiation damage
- “Native” = dataset after radiation damage

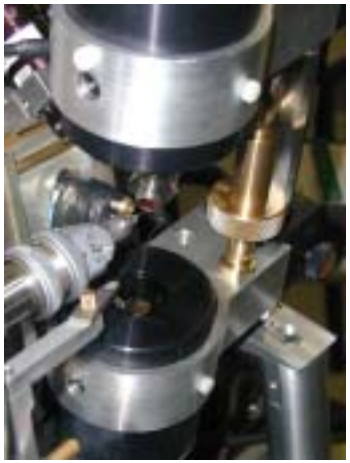
  – ===> Radiation damage Induced Phasing (RIP)

Data Collection Strategies

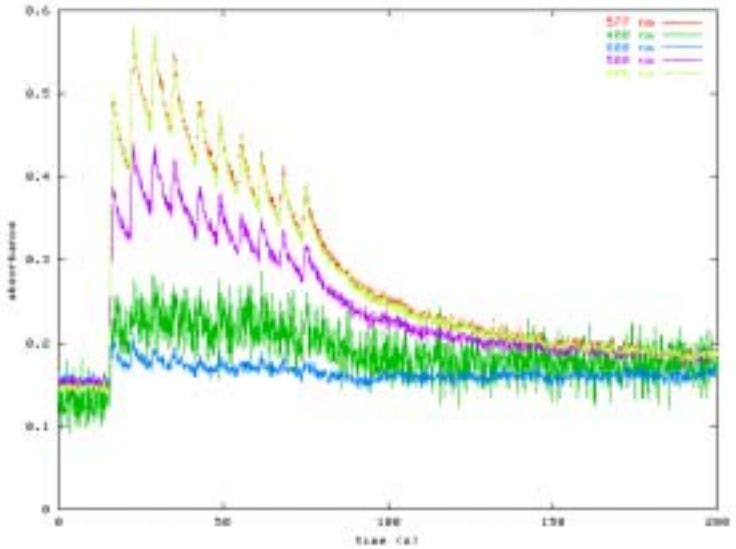
- Collect “before” dataset, “burn”, collect “after” dataset.
  - Can optimize burn -- not too much, not too little
  - Can use existing software in SIR mode (and SHELXC in RIP mode)
- Collect multiple datasets until crystal burns out
  - No need for burn optimization
  - In principle can yield accurate zero and final dose extrapolated data
  - Longer data collection necessary
  - Need a method of selecting which datasets

A-burn-B

- Done with no attenuation
- Burn optimization
  - Burn length by
    - Trial and error
    - Spectrophotometric monitoring of radical formation



Ravelli, Murray, Weik, Theveneau, Garman



RIP datasets

Protein	Resolution (Å)	SG	% solvent	Site	Sites:#residues
Insulin	1.4	$I2_13$	64	SS	3:51
Trypsin	1.4	$P2_12_12_1$	41	SS	3:229
RNAse	1.65	$P3_221$	50	SS	3:102
HEWL	1.6	$P4_32_12$	35	SS	4:129
RNA	1.4	$P2_12_12_1$	34	Br	8:48
PYP	1.4	$P6_5$	35	S-C	1:122
DNA	1.4	$P3_121$	50	Br	8:83
Thaumatin	1.6	$P4_12_12$	56	SS	8:207
A	1.7	$P2_12_12_1$	59	SS	1:500
Fibronectin	1.6	$P4_12_12$	30	SS	4:90
C	1.2	C2	28	CO	0:82

Successes in RIP

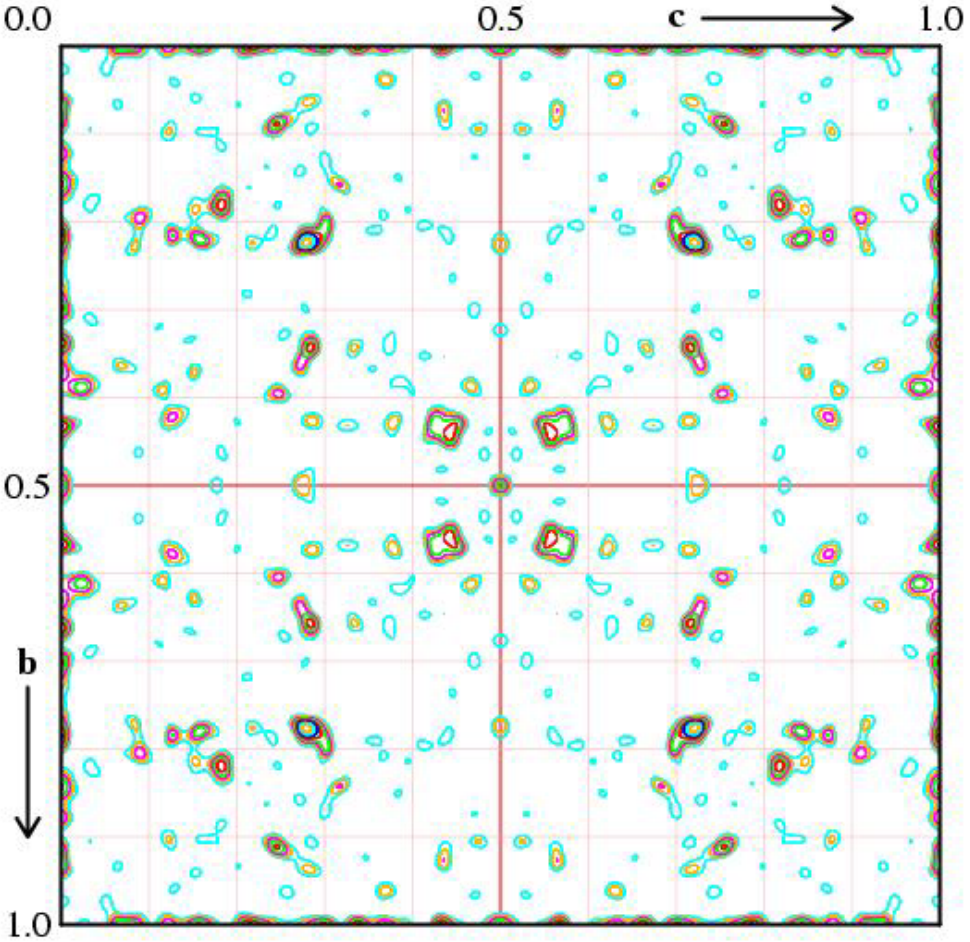
- Generally high resolution data
- Generally with SS
- One by just S
- None by just carboxylates

Solving structures by RIP

- SHARP
  - Explicitly treated
  - Accurate phases, but speed is an issue
- SHELX
  - Preparation of DeltaFs (F_{before}-F_{after})
    - SHELXC
    - XPREP
  - SHELXD (best CC's for good data ~30%)
  - SHELXE

RIP Harker Sections

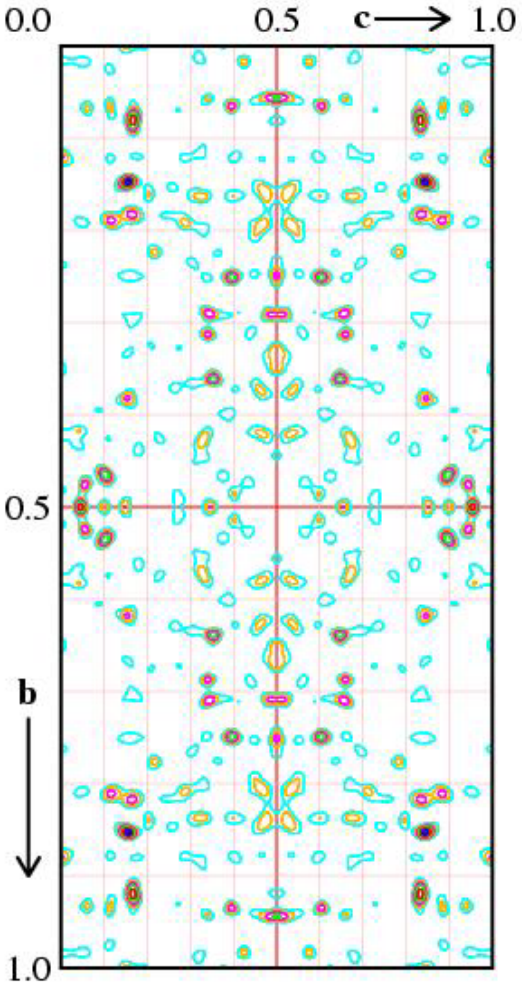
Patterson section X = 0.0000 for xp_1_0.98850.hkl
Space group: I2(1)3
Cell: 78.320 78.320 78.320 90.000 90.000 90.000
+Y down, +Z across, 512 x 512 grid, contour interval = 1.0 sigma
Super-sharpened, origin removed



Insulin

RIP Harker Sections

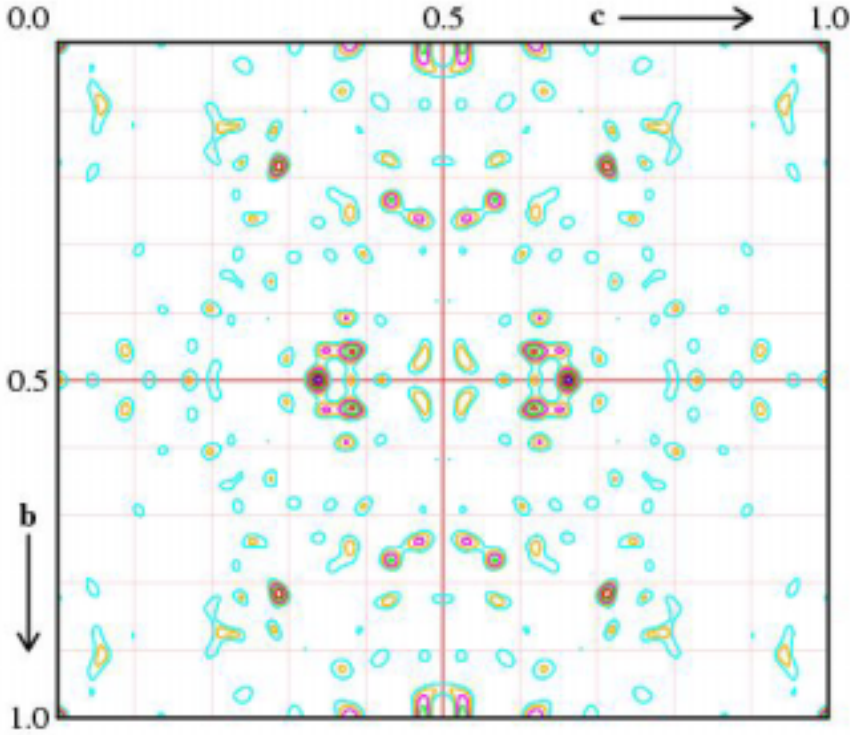
Patterson section X = 0.5000 for xp_1_0.98188.hkl
Space group: P4(3)2(1)2
Cell: 78.770 78.770 36.820 90.000 90.000 90.000
+Y down, +Z across, 512 x 256 grid, contour interval = 1.0 sigma
Super-sharpened, origin removed



Lysozyme

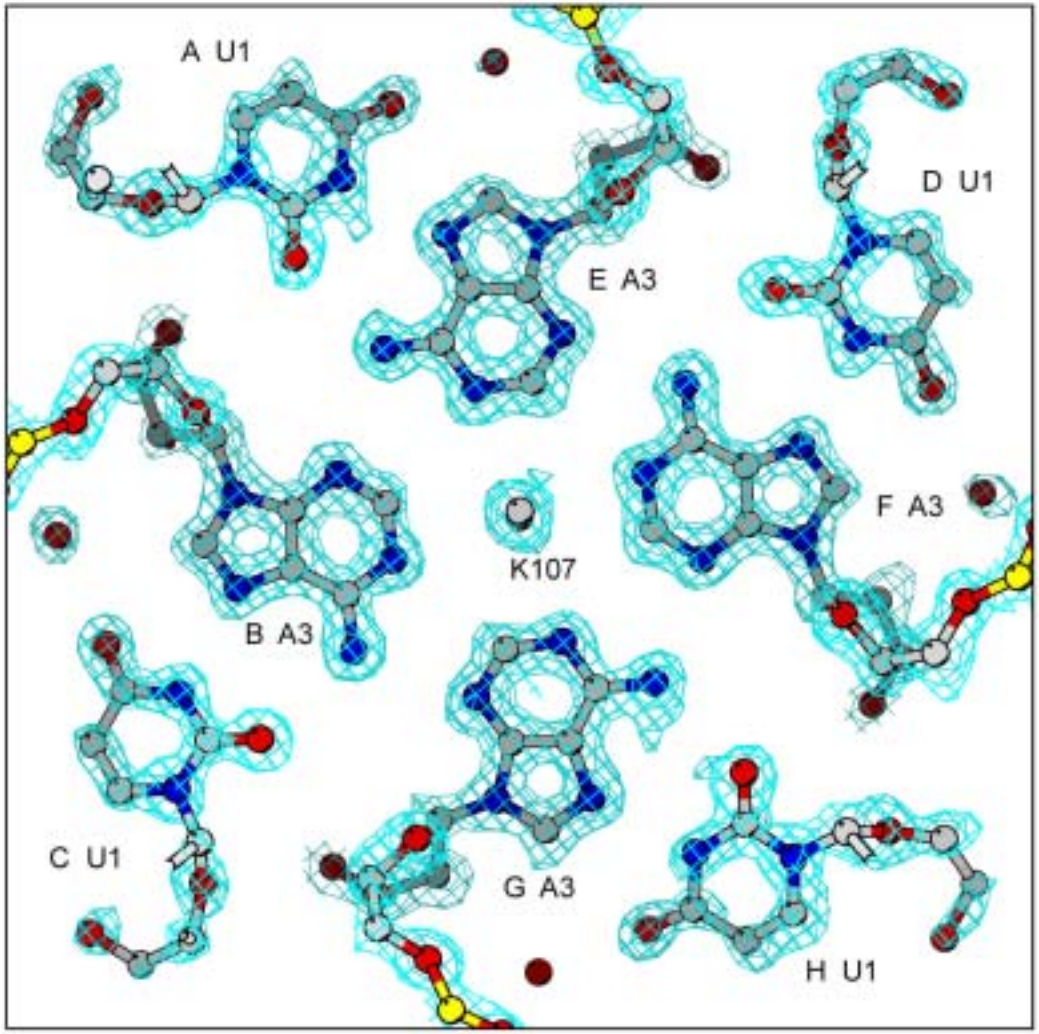
RIP Harker Sections

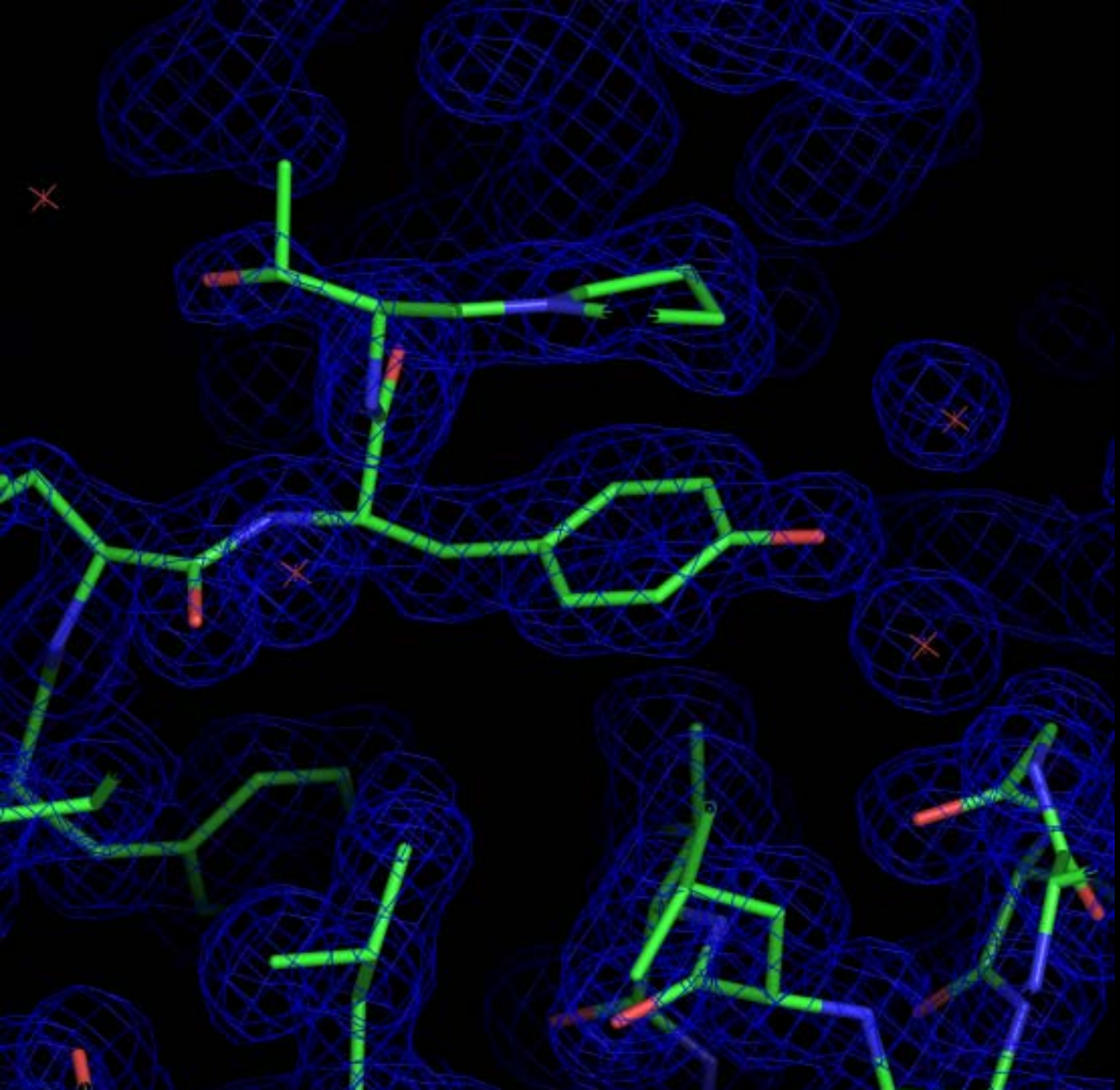
Patterson section X = 0.5000 for xp_1_0.99525.hkl
Space group: P2(1)2(1)2(1)
Cell: 54.253 58.332 66.479 90.000 90.000 90.000
+Y down, +Z across, 256 x 256 grid, contour interval = 1.0 sigma
Super-sharpened, origin removed



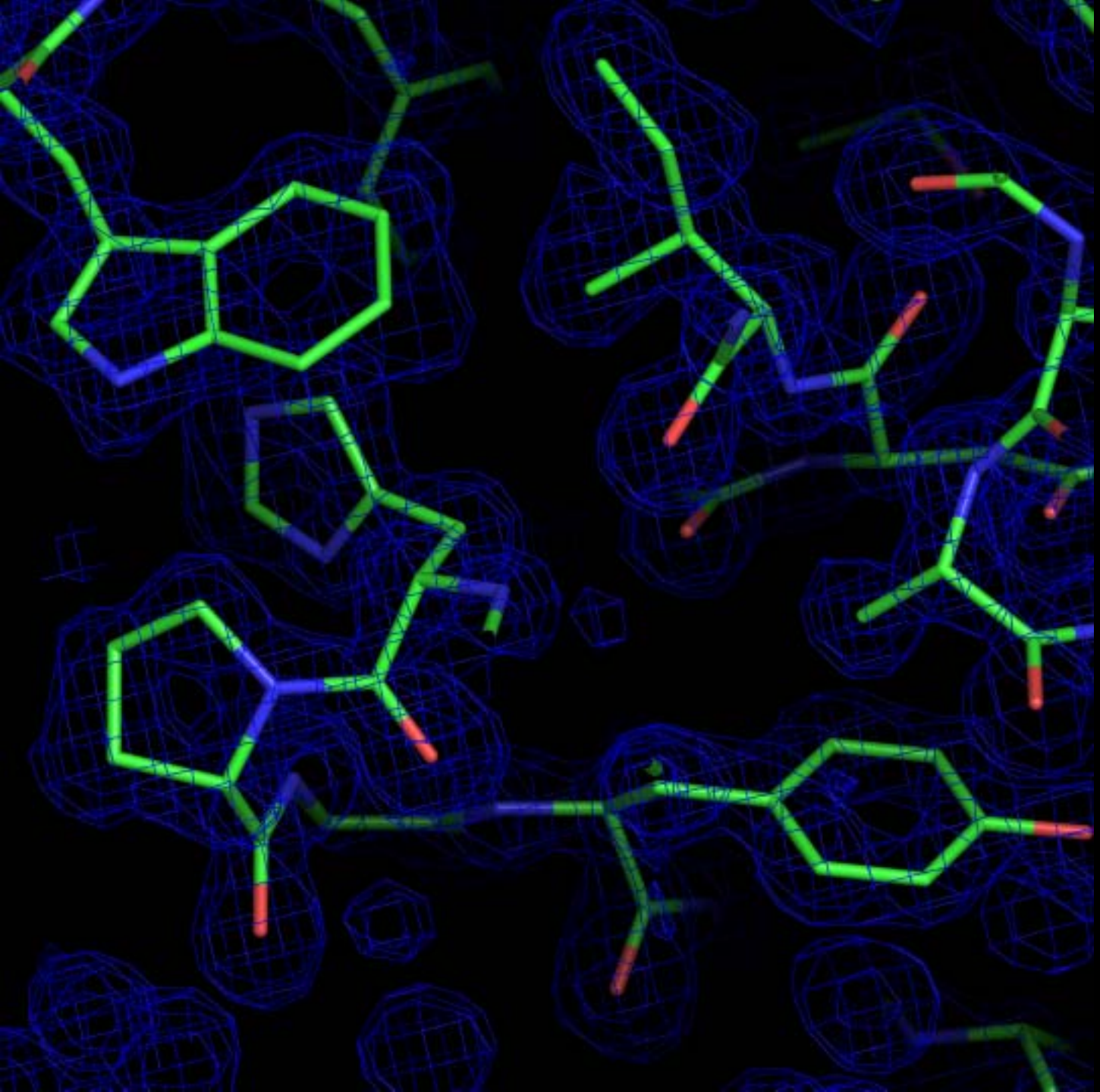
Trypsin

DNA/RNA Hybrid





Insulin
SHELXE
Map
1.0 σ



Trypsin
SHELXE
Map
1.0 σ

RIP issues

- Finding sites can be improved
- Can scaling for RIP be improved?
- Can extracting phases can be improved?

Finding Sites

- F_{before} - F_{after} Pattersons are different than SIR pattersons
  - "Positive" peaks arise like SIR, MIR etc (existing positions disappear)
  - "Negative" peaks from SH groups swinging away (new positions appear)

Improving Scaling

- Possible to improve scaling?
  - Many short negative vectors in F_{before} - F_{after} Pattersons
  - Many negative peaks (= electrons gained during the experiment) in unexpected and unlikely positions.
- Apply artificial scale to F_{after}
  - Observe effects on peak height, noise

Future Prospects

- A-Burn-B method
  - Improve site identification
  - Improve phase accuracy
  - Improve burning method
    - Maximize specific changes
    - Minimize general changes
- Multiple dataset method/Zero dose extrapolation
  - Bricogne -SHARP
  - Weiss and Warkentin Acta Cryst. (2004). D 60 , 686-695
  - Bodek and Otwinoswski --SCALEPACK
  - Kabsch -XSCALE
  - Diederichs and Ravelli Acta Cryst. (2003). D 59 , 903-90
- Generally
  - Our understanding of the effects of radiation damage is incomplete
  - Can we use low resolution differences?

RIP advantages

- No modification of the protein necessary
- Can be performed on a fixed wavelength beamline
- Fast data collection

Cubic Insulin		
Collect 45 frames attenuated	5 min	
Burn crystal 30 s unattenuated		1 min
Collect 45 frames attenuated	5 min	
Run xprep and shelxd to find 3 sites	1 min	
Cycle shelxe -b to find 40 sites		3 min
Total (excluding arp/warp)		15 min

- 2-for-1: Understanding radiation damage
  - Improves data quality
  - Can give useful phase information

Acknowledgements

- Raimond Ravelli
- George Sheldrick
- Elspeth Garman